



OPEN

## The structure, function and expression analysis of the nodulin 26-like intrinsic protein subfamily of plant aquaporins in tomato

Yuxiang Zhang, Shihong Fei, Yunmin Xu, Yong He, Zhujun Zhu<sup>✉</sup> & Yuanyuan Liu<sup>✉</sup>

The nodulin 26-like intrinsic protein (NIP) family belonging to a group of aquaporin proteins is unique to plants. NIPs have a wide of transport activities and are involved in developmental processes and stress tolerance. The well reported Lsi1 and Lsi6 belonging to NIP III were characterized as Si transporters. However, except Lsi1 and Lsi6, most NIPs remain unknown. Here, we identified 43 putative aquaporins in tomato. We found there are 12 NIPs, including 8 NIP I proteins, 3 NIP II proteins, and 1 NIP III protein among the 43 aquaporins. Also, there are two Si efflux transporters SLsi2-1 and SLsi2-2 identified by using Lsi2 proteins from other species. By analysing the phylogenetic relationships, conserved residues and expression patterns, we propose that three NIP I members (SINIP-2, SINIP-3 and SINIP-11) may transport water, ammonia, urea, and boric acid, and contribute to pollen development. Three NIP II proteins (SINIP-7, SINIP-9 and SINIP-12) may be boric acid facilitators, and affect plant growth and anther development. Overall, the study provides valuable candidates of Si transporters and other NIP proteins to further explore their roles in uptake and transport for silicon, boron, and other substrates in tomato.

Silicon (Si) is a common element, and its abundance in the earth's crust is second only to oxygen, reaching 28.8%<sup>1</sup>. Si is an essential element for animals and is closely related to animal life activities<sup>2</sup>. Although Si is not a vital element for plants<sup>1,3</sup>, a large number of studies have shown that Si has beneficial effects on plant growth and development, especially under stressed conditions, such as drought, salt, heavy metal, high temperature, radiation damage, freezing, nutrient imbalance<sup>4–6</sup>. In plants, Si has been proposed as a quasi-essential element, and the International Plant Nutrition Institute (IPNI) defined Si as a "beneficial substance" in 2015. The main forms of Si existing in soil are silica (SiO<sub>2</sub>) and silicate, and only a small part exists in the form of silicic acid Si(OH)<sub>4</sub><sup>7</sup>, which is the only form of Si that can be absorbed by plants<sup>8</sup>.

Although Si is found in most plants<sup>7</sup>, the content of Si varies greatly among different plants<sup>9</sup>. According to the differences in Si absorption modes, plants were classified into active, passive and rejective Si absorption plants<sup>10</sup>. In addition, plants are divided into three types: high Si accumulators, intermediate Si accumulators and low Si accumulators<sup>3</sup>. Monocots such as rice and wheat belong to high Si accumulation plants and active Si absorption plants, and their Si content can reach 1.0–10.0% of dry weight. Most of the eudicots, such as cucumbers, belong to intermediary Si accumulation plants and passive Si absorption plants, and the Si content in their bodies is about 0.2–1.0% of dry weight. Some of the eudicots, such as tomato, belong to low Si accumulation plants and Si rejected plants, and their Si content is only less than 0.2% of dry weight<sup>11</sup>.

The absorption of Si by plants is achieved mainly by two steps. Firstly, silicic acid is absorbed by roots, and then silicic acid is transported to aboveground by transpiration pull<sup>12</sup>. Among them, although transpiration pull can promote the transfer of Si in plants from roots to shoots, studies have proved that transpiration pull is not the reason for the difference of Si accumulation in different plants, but the uptake ability of root plays a major role in Si uptake and transport in plants<sup>12</sup>. The uptake process of Si is mediated by specific transmembrane proteins, Low silicon rice 1 (Lsi1), identified as Si transporters. The first identified Si transporter in higher plants is OsLsi1

Collaborative Innovation Center for Efficient and Green Production of Agriculture in Mountainous Areas of Zhejiang Province, College of Horticulture Science, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China. ✉email: zhuzj@zafu.edu.cn; yylu@zafu.edu.cn

in rice<sup>13</sup>. OsLsi1 is located in the distal side of the plasma membrane of the exodermis and endodermis cells. Its function is to facilitate Si uptake into cells of exodermis from soil solution, and promote Si uptake into cells of endodermis from aerenchyma<sup>13</sup>. Shortly after the discovery of OsLsi1, a second Si transporter gene, OsLsi2, was uncovered<sup>14</sup>. OsLsi2 is located in the proximal side of the plasma membrane of the exodermis and endodermis cells. Its function is to facilitate Si efflux into aerenchyma from exodermis cells and facilitate Si efflux into soil from endodermis cells<sup>14</sup>. Obviously, OsLsi1 is an influx transporter, while OsLsi2 is an efflux transporter<sup>13,14</sup>. A third Si transporter identified as influx transporter is OsLsi6 in rice<sup>15</sup>. OsLsi6 is primarily located in the parenchyma cells of xylem in the leaf sheaths and leaves, which is responsible for xylem unloading and further distribution of Si in aerial tissues<sup>15,16</sup>.

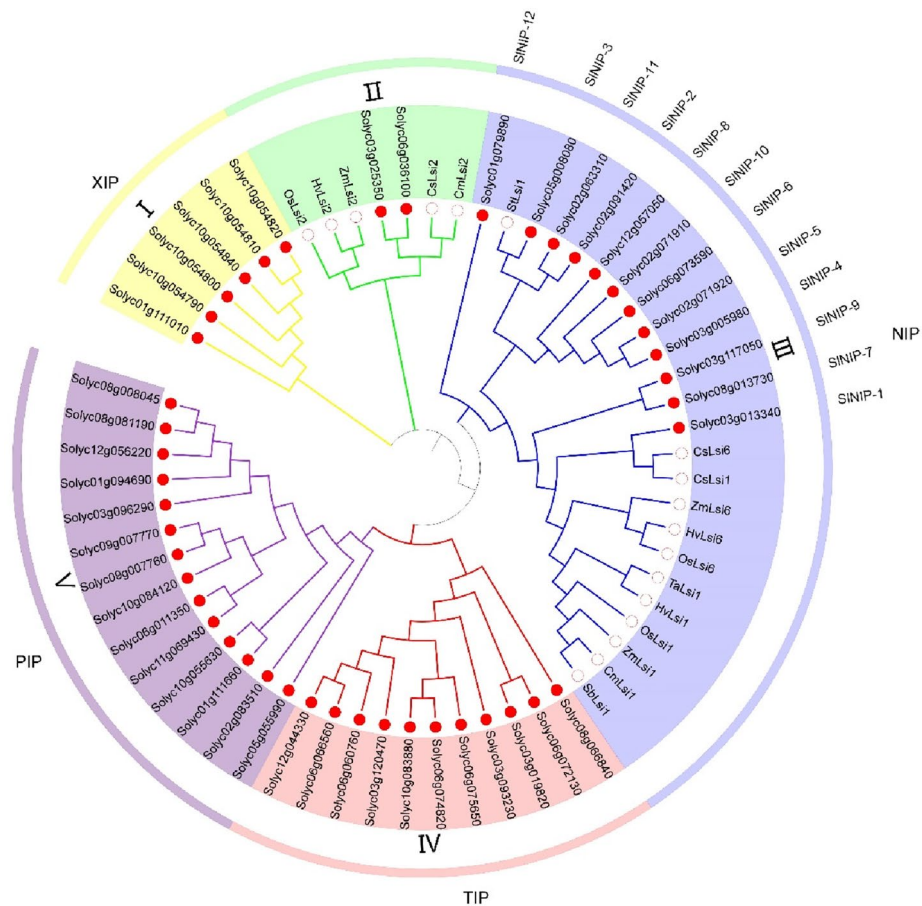
It has been reported that Lsi1 and Lsi6 belong to the Nodulin 26-like intrinsic proteins (NIPs) subfamily of the aquaporin (AQP) family<sup>13</sup>, while Lsi2 falls into the ion transporter superfamily<sup>14</sup>. AQPs are water channel proteins, also called major intrinsic proteins (MIPs) that facilitate the transport of water and a variety of solutes, such as boric acid, NH<sub>4</sub><sup>+</sup>, glycerol or urea<sup>17</sup>. Based on the sequence similarity and the localization, plant AQPs can be divided into five groups. Except NIPs, the other four groups include the plasma membrane intrinsic proteins (PIPs), the tonoplast intrinsic proteins (TIPs), the small basic intrinsic proteins (SIPs), and the X intrinsic proteins (XIPs)<sup>18–21</sup>. Among these AQPs, NIPs are unique to plants. Furthermore, NIPs can be classified into three subgroups—NIP I, NIP II and NIP III—with respect to the composition of the aromatic/arginine (ar/R) selectivity filter which is formed by four residues from helices 2 (H2), H5, loop E (LE1) and LE2<sup>22,23</sup>. The size and chemical properties of the residues in ar/R selectivity filter may determine the size and chemical properties of substrates NIPs can transport. NIP I can transport water, glycerol<sup>24</sup> and lactic acid<sup>25</sup>; NIP II can transport boron<sup>26</sup> and urea<sup>27</sup>; whereas, NIP III have transport activities for silicon, arsenic and selenium<sup>22</sup>. Indeed, unlike AQP1 belonging to NIP I which contains two bulky amino acid residues in the H2/H5, OsLsi1 belonging to NIP III contains two small residues has a broader pore diameter of the channel<sup>28</sup>. Therefore, OsLsi1 can transport relatively larger molecules of silicic acid (4.38 Å), while the substitution of larger Phe and His for smaller Gly and Ser at H2 and H5 constrains the AQP1 pore to 2.8 Å permitting the flux of water (diameter 2.4 Å) and the exclusion of bulkier solutes<sup>23,28</sup>.

Si is extremely low in tomato plant. At present, there is still a lack of research on the Si uptake and transport in tomato, and the reasons for its low Si accumulation are still controversial. One study first reported that SLsi1 with 109-AA instead of 108-AA spacing between NPA domains showed loss of Si permeability, compared with the functional Lsi1 transporters in high Si-accumulators<sup>29</sup>. The other study believed that SLsi1 is normally functional, but loss-of-function of SLsi2 cannot transfer silicic acid absorbed from roots to xylem, thus resulting in low Si accumulation in shoot<sup>30</sup>. Although tomato is a Si excluder, Si could benefit tomato a lot on many aspects, such as seed germination, shoot growth, the root/shoot ratio, osmotic adjustment, photosynthetic rate, hydraulic conductance, transpiration, antioxidant systems<sup>31</sup>. Evidences on Si transporters in tomato, even in eudicots, are relatively insufficient. The functions of other NIP proteins other than SLsi1 in tomato are also still unknown. In the present study, Lsi1 and Lsi6 Si influx transporters and Lsi2 Si efflux transporters from other species were set to blast tomato homologues. 43 aquaporin proteins homologous to Lsi1 and Lsi6, and 2 ion transporter proteins homologous to Lsi2 were identified in tomato. Detailed information, including their phylogenetic relationships, gene structures, conserved domains, subcellular localizations, tissue expression profiles, and cis-acting elements were analyzed. The studies will provide a solid basis for further functional characterization of tomato Si transporters as well as other NIP members, and exploring molecular mechanism of their uptake and transport for silicon, boron, and other substrates in tomato.

## Results

**Identification of Lsi1/Lsi2/Lsi6 homologues of aquaporin and ion subfamily in tomato.** Tomato Si transporter homologues were identified from the tomato genome database (SGN, <https://solgenomics.net>) by a BLAST search using the full-length protein sequences of reported Si transporters Lsi1/Lsi2/Lsi6 from rice (*Oryza sativa* L.) and cucumber (*Cucumis sativus* L.). The amino acid sequences of Lsi1/Lsi2/Lsi6 from rice and cucumber were downloaded from NCBI database (Supplementary Table S1). A total of 43 tomato aquaporin proteins and 2 ion transporter proteins were obtained after deletion of repetitive sequences.

**Phylogenetic analysis and classification of tomato Lsi1/Lsi2/Lsi6 homologues.** To reveal the phylogenetic relationships of 45 tomato homologues, the amino acid sequences of reported 17 Lsi1/Lsi2/Lsi6 Si transporters from other species were included (Supplementary Table S1) for the construction of phylogenetic tree. These 62 proteins can be classified into five clades (I–V, Fig. 1). Notably, each clade contains tomato homologues, six are assigned to Clade I. Two are classified into Clade II that includes reported Lsi2 Si transporters. Twelve are members of Clade III, together with the well reported Lsi1/Lsi6 Si transporters. Eleven are divided into Clade IV, and fourteen belong to Clade V. Actually, members in Clade I, Clade III, Clade IV and Clade V correspond to members in XIPs, NIPs, TIPs, and PIPs described in current literature<sup>32</sup>. Twelve tomato homologues were grouped in Clade III with StLsi1 of potato, CsLsi1 and CsLsi6 of cucumber, ZmLsi1 and ZmLsi6 of maize, HvLsi1 and HvLsi6 of barley, OsLsi1 and OsLsi6 of rice, TaLsi1 of wheat, CmLsi1 of pumpkin, SbLsi1 of sorghum. It has been reported that OsLsi1<sup>13</sup>, OsLsi6<sup>15</sup>, HvLsi1<sup>33</sup>, ZmLsi1 and ZmLsi6<sup>34</sup>, CsLsi1<sup>35</sup> belong to NIP III subfamily of aquaporins, suggesting the twelve tomato homologues in Clade III also belong to the NIP subgroup. We name them as SINIP-1 (Solyc03g013340), SINIP-2 (Solyc02g091420), SINIP-3 (Solyc05g008080), SINIP-4 (Solyc03g005980), SINIP-5 (Solyc02g071920), SINIP-6 (Solyc06g073590), SINIP-7 (Solyc08g013730), SINIP-8 (Solyc12g057050), SINIP-9 (Solyc03g117050), SINIP-10 (Solyc02g071910), SINIP-11 (Solyc02g063310) and SINIP-12 (Solyc01g079890), respectively. Solyc03g025350 and Solyc06g036100 clustered in Clade II with OsLsi2 of rice, HvLsi2 of barley, ZmLsi2 of maize, CsLsi2 of cucumber and CmLsi2 of pumpkin. Based on pro-

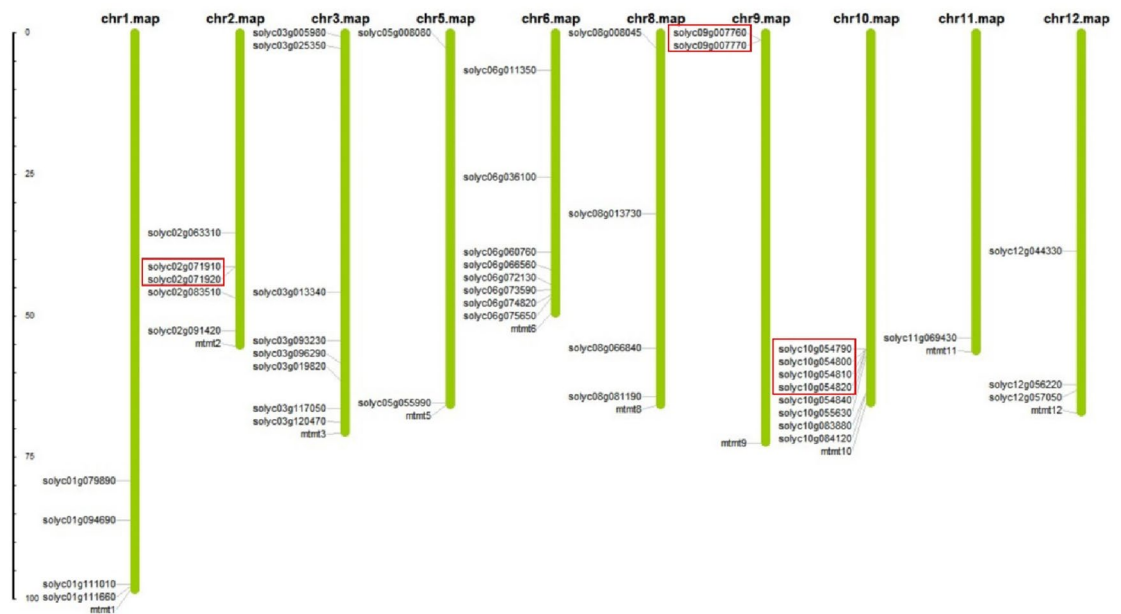


**Figure 1.** Phylogenetic analysis of 45 homologous proteins of Si transporters identified in tomato and Si transporters available in other species. The tree was constructed using the maximum likelihood method (MLM) with 1000 bootstraps. Each of the five different clades is represented by a specific color and named as I, II, III, IV and V. Proteins from tomato are marked with red solid circles. Proteins from other species are marked with hollow circles. *Hv* stands for barley (*Hordeum vulgare* L.), *Ta* stands for wheat (*Triticum aestivum* L.), *Zm* stands for corn (*Zea mays* L.), *Sb* stands for sorghum (*Sorghum bicolor* L.), *Cm* stands for pumpkin (*Cucurbita moschata* L.), *St* stands for potato (*Solanum tuberosum* L.), *Os* stands for rice (*Oryza sativa* L.), *Cs* stands for cucumber (*Cucumis sativus* L.).

tein sequence similarity and their relationship in the phylogenetic tree, the two orthologous Lsi2-like proteins of tomato in Clade II were named as SILsi2-1 (Solyc06g036100) and SILsi2-2 (Solyc03g025350) (Supplementary Table S2).

Lsi1 and Lsi6 in other species have been identified as Si influx transporters, and Lsi2 has been characterized as Si efflux transporters. In Lsi1/Lsi6 proteins-containing Clade III, the orthologous SINIP-1 was the most homologous to these Si influx transporters, and reasonably predicted to have Si influx transport activity. Similarly, the orthologous SILsi2-1 and SILsi2-2 were proposed to have efflux transport activity for Si. However, in a previous study, no Si transport activity of SILsi1 (designated SINIP-1 in this study) was detected in *Xenopus oocytes*<sup>29</sup>. This report is not consistent with the other study which determined that SILsi1 was a functional influx transporter<sup>30</sup>. SILsi2-L1 (designated SILsi2-1 in this study) and SILsi2-L2 (designated SILsi2-2 in this study) were reported no efflux transport activity for Si<sup>30</sup>. However, the function of aquaporins in tomato remain to be largely uncharacterized, except SINIP-1, members in Clade III that belong to NIP subgroup are still unknown.

**The chromosomal localization of 45 Lsi1/Lsi2/Lsi6 homologous genes in tomato.** The chromosomal localizations of 45 tomato Lsi1/Lsi2/Lsi6 homologues were determined to visualize their genomic position information. They were distributed on 10 of 12 chromosomes except for chromosome 4 and chromosome 7 (Fig. 2). Among them, eight genes were mapped to chromosome 3, 6 and 10; two on chromosome 5 and 9, three on chromosome 12, four on chromosome 1 and 8, five on chromosome 2; there was only one gene on chromosome 11 (Fig. 2). It should be noted that gene duplication may occur among these tomato Si transporter homologues, since the loci Solyc10g054790, Solyc10g054800, Solyc10g054810 to Solyc10g054820 are found next to each other on chromosome 10. Also the gene pairs Solyc02g071910/Solyc02g071920, Solyc09g007760/Solyc09g007770 were found on chromosome 2 and chromosome 9, respectively. Interestingly, Solyc10g054790,



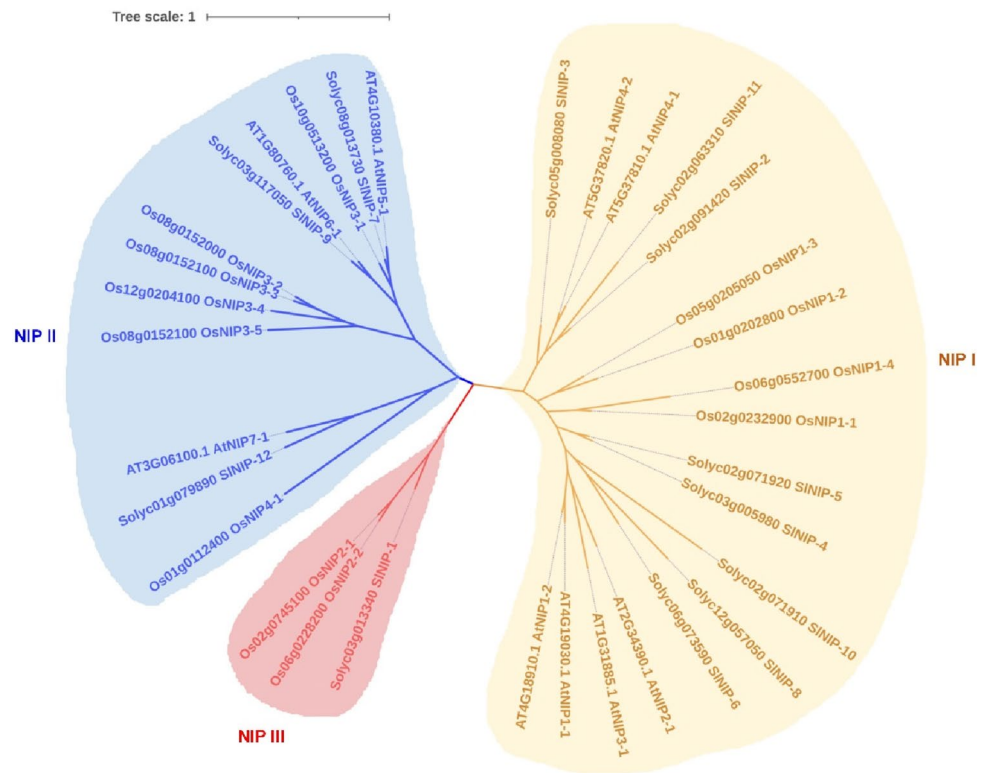
**Figure 2.** Distribution of homologous genes of Si transporters on tomato linkage groups (LGs). The 45 homologous genes of Si transporters were mapped onto 10 LGs in the tomato genome. The LG number was shown on the top of each LG. The rule on the left indicates the physical map genome. The LG number was shown on the top of each LG. The rule on the left indicates the physical map distance among genes (Mbp). The red boxes indicate tandem array genes.

Solyc10g054800, Solyc10g054810 and Solyc10g054820 were clustered together in Clade I, Solyc02g071910/Solyc02g071920 were grouped together in Clade III, and Solyc09g007760/Solyc09g007770 were clustered together in Clade V (Fig. 1), which showed high similarity based on protein sequences, indicating they may share common functions with each other.

**Analysis of exon–intron structure.** The exon and intron information of 45 tomato Lsi1/Lsi2/Lsi6 homologues were searched in SGN database and sorted by GSDS. The color bar on the left side of the gene name represents the clades of these genes in the evolutionary tree. Exon–intron analysis showed that the size and the number of the exons are highly conserved within each clade, but significantly different among the clades. Most members of the 45 tomato Lsi1/Lsi2/Lsi6 homologous genes contained three to five exons, but a few contained less than three or more than five exons. Members of Clade I and Clade II are characterized by three exons. Most members of Clade III have five exons. The majority of the members of Clade IV features three exons, while most members of Clade V contain four exons (Supplementary Fig. S1A,B).

**Gene structure and motif composition analysis.** Motifs are highly conserved amino acid residues in homologous proteins, which may play important roles in the structure and function of active proteins. The conserved motifs of these 45 tomato Lsi1/Lsi2/Lsi6 homologues were analyzed by MEME. Ten conserved motifs were identified (Supplementary Fig. S1C). Among them, Motif 9 only exists in the members of Clade II and this clade does not contain any other motifs, which might contribute to the functional divergence of Lsi2 proteins in Clade II. It is reasonable Lsi2 was ion transporter protein, differing from NIP proteins. Motifs 1, 2 and 3 were found in Clade III that clustered by Lsi1/Lsi6 proteins, which were also shared in Clade I, IV and V. Interestingly, members of Clade I and V, and some members of Clade IV, but not any members of Clade III, also shared Motif 7. Motif 8 exists in all members of Clade V and some members of Clade I; in addition, members in Clade I solely contain Motifs 5, 6 and 10, and members in Clade V uniquely possess Motif 4 (Supplementary Fig. S1A,C). These observations suggest genes in Clade I, IV and V may have evolved from genes in Clade III, sharing common motifs (Motifs 1, 2 and 3) and similar functions with members in Clade III, and further gained some additional motifs (such as Motif 7) and corresponding functional diversification afterwards.

**Evolutionary analysis of NIP subfamily from tomato, Arabidopsis and rice.** The well-characterized Lsi1 channels belong to the NIP subfamily of aquaporins. In our study, we have found 12 members that in Clade III belong to the NIP subgroup (Fig. 1). However, except SINIP-1, the other members in NIP subgroup are still unknown. Therefore, we have made an in-depth study on these twelve NIP proteins. To well understand their role in substrate selectivity, we constructed phylogenetic analyses of these 12 NIPs with orthologues from *Arabidopsis*<sup>36</sup> and rice<sup>37</sup> (Fig. 3), the accession numbers of the corresponding proteins were listed in Supplementary Table S3. As phylogenetic tree showed that the *Arabidopsis* and rice NIPs were distributed in two (I and II) and three (I, II and III) subgroups, respectively, which is well consistent with previous reports<sup>29</sup>. The 12 tomato NIPs fall into three NIP subgroups (Fig. 3). SINIP-1 is grouped with well-defined OsNIP2-1 (OsLsi1) and OsNIP2-2 (OsLsi6) that belong to NIP III subgroup. SINIP-7, SINIP-9 and SINIP-12 fall into NIP II sub-



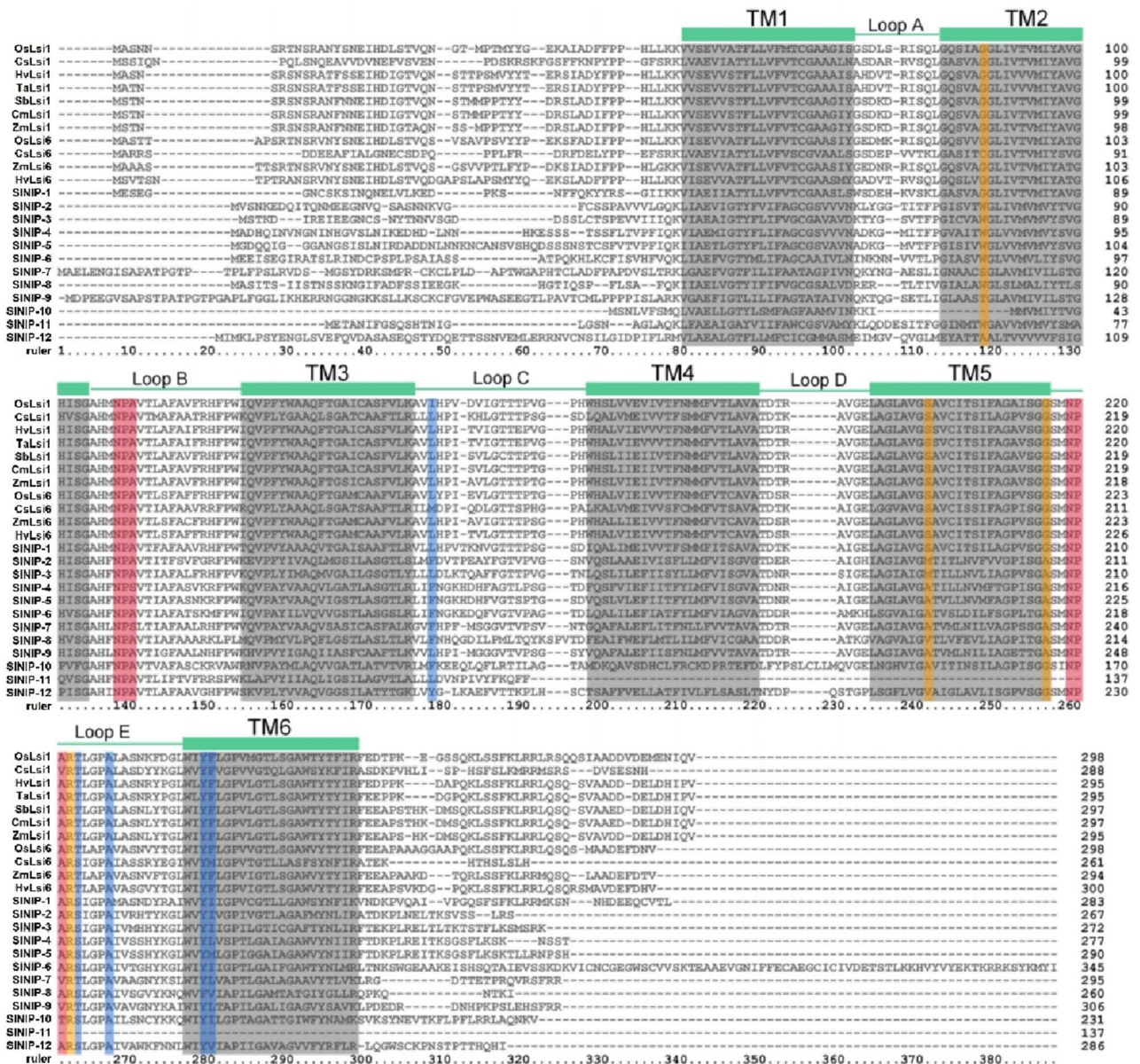
**Figure 3.** Evolutionary analysis of twelve, nine, twelve NIP proteins from tomato, *Arabidopsis* and rice. The phylogenetic tree was constructed by MEGAX using maximum likelihood method, and was modified by iTOL. Members in NIP I, II and III were shown by red, blue and yellow colors, respectively.

group, and were tightly clustered together with AtNIP5-1, AtNIP6-1 and AtNIP7-1, respectively. Eight out of the twelve tomato NIPs belong to NIP I, including SINIP-2, SINIP-3, SINIP-4, SINIP-5, SINIP-6, SINIP-8, SINIP-10 and SINIP-11. Among them, SINIP-2, SINIP-3 and SINIP-11 showed relatively closed to AtNIP4-1 and AtNIP4-2, whereas SINIP-4, SINIP-5, SINIP-6, SINIP-8 and SINIP-10 showed highly closed to AtNIP1-1, AtNIP1-2, AtNIP2-1, and AtNIP3-1 (Fig. 3).

**Prediction of physicochemical properties and identification of conserved domain.** It is well known that NIPs of aquaporin subfamily were characterized for six transmembrane domains (TMDs) connected by five loops (loop A—loop E)<sup>38</sup>, two highly conserved NPA (asparagine-proline-alanine) motifs on loop B and loop E, an ar/R selectivity filter and Froger's positions<sup>39</sup>. Sequence alignments of twelve tomato NIPs were performed with ClustalX2 (Fig. 4). The TMDs of each protein were predicted by TMHMM2.0 online tool, the positions of TMDs are marked with gray areas, and the numbers of TMDs of each protein are recorded in Table 1. All the identified Si transporters of Lsi1/Lsi6 in other species contain six TMDs, and most of the twelve tomato NIPs also contain six TMDs. However, there is an exception for SINIP-10 and SINIP-11, which possess four and three TMDs, respectively. The second TMD in SINIP-10 is incomplete, while SINIP-11 missed the fourth, fifth and sixth TMDs (Table 1).

For the members of aquaporin family, NPA domains will generate electrostatic repulsion to protons, and then form a channel, while ar/R selectivity filter and Froger's positions will affect the specific substrate of the channel. All Si transporters identified in other species have two NPA domains, only with an exception for CsLsi1, which have NPV on Loop E with the alanine (A) in the third position replaced by threonine (T). However, NPA motif changed into NPV does not disturb Si transport activity of CsLsi1<sup>35</sup>. As for tomato, 11 of the 12 tomato NIPs all carried two NPA/S/V/T motifs, except for SINIP-11 only having one NPA motif on Loop B. A specific length of 108 AA between the two NPA domains is a necessary and selectivity feature for Si among all Si-transporting plants<sup>29</sup>. These already validated Si-transporting Lsi1 proteins all presented 108 AA residues, such as CsLsi1, HvLsi1, TaLsi1, SbLsi1, CmLsi1, ZmLsi1 and OsLsi1. Whereas most of the tomato NIPs have 109 residues in the spacing, and only SINIP-7 and SINIP-9 have 108 AA. Moreover, SINIP-8 has 112 AA and SINIP-10 has 115 AA between the two NPA domains. However, the spacing itself is unlikely to be required for Si permeability. Supporting evidences showed that SiLsi1/SINIP-1 possesses 109 AA in the spacing and shows transport activity for Si<sup>30</sup>.

For most members of NIP III subgroup from various species, including SINIP-1 from tomato, four amino acids from helix 2 (H2), helix 5 (H5), and loop E (LE1 and LE2) that constitute ar/R selectivity filter are GSGR, except for CsLsi6, where a small G in the H2 is substituted with the bulkier C residue (Table 1). However, the residue at the H2 position is not critical for Si transport. For example, when glycine (G) was substituted by alanine



**Figure 4.** Identification of conserved domains of Si transporter Lsi1/6 homologues (SINIP-1 to SINIP-12) in tomato and Si transporters Lsi1/6 reported in other plants. Protein sequence alignment of functionally known Si transporters Lsi1/6 in other species and the newly identified homologues in tomato. The gray region represents the transmembrane (TM) domains, the red region represents the conserved NPA domains, the yellow region represents the ar/R selectivity filter sites, and the blue region represents the Fröger's residue sites.

(A) at H2 of OsLsi1, the transport activities for Si were unaffected<sup>40</sup>. Compared to residue at H2, residue at H5 is required for Si transport activity. When serine (S) at the H5 position was substituted by isoleucine (I), and the transport activity of OsLsi1 for Si was totally lost<sup>40</sup>. In addition, when both residues at the H2 and H5 were substituted, Si transport activities of OsLsi1 as well as transport activities for other solutes are completely lost. Most of tomato NIPs belong to NIP I subgroup possessing an ar/R filter consisting of W, V, A and R, except SINIP-6 with ar/R selectivity filters of WIAR, and SINIP-10 and SINIP-11 whose ar/R selectivity filters are incomplete. For SINIP-7, SINIP-9 and SINIP-12, which belong to the NIP II subgroup, their ar/R selectivity filters are SIAR, TIAR, and AVGR, respectively. It proposed us to speculate that SINIP-6, SINIP-7 and SINIP-9 with I at H5, and SINIP-10 and SINIP-11 with incomplete ar/R selectivity filter might have no transport activities for Si. It is worth testing the transport activities of SINIP-1, SINIP-2, SINIP-3, SINIP-4, SINIP-5, SINIP-8 with ar/R selectivity filter consisting of WVAR and SINIP-12 with ar/R selectivity filter consisting of AVGR.

Fröger's positions are composed of five amino acids, which are located in highly conserved regions. Position 1, located in the terminal part of the third transmembrane segment, Positions 2 and 3 are located in loop E, and Positions 4 and 5 are located in the sixth transmembrane segment<sup>39</sup>, Fig. 4). Residues in Position 3 (A) and Position 4 (Y) are conserved in almost all identified Si transporters. Position 2 is defined as a polar residue

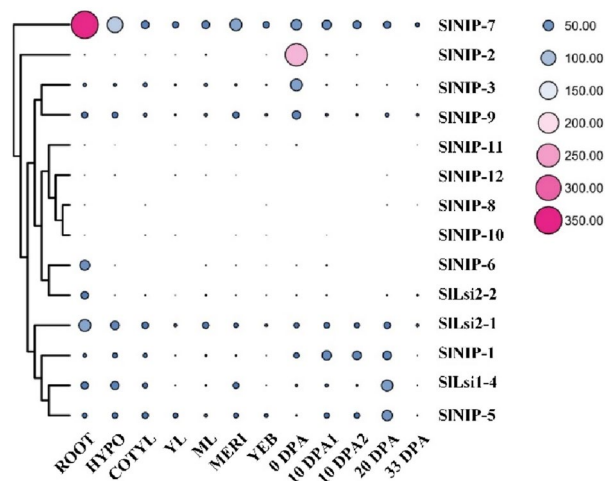
Name	Len	MW	pI	GR	TM	ar/R selectivity filter				Froger residue					NPA motif			LOC
						H2	H5	LE1	LE2	P1	P2	P3	P4	P5	LB	LE	DB	
SINIP-1	283	30.08	8.76	0.37	6	G	S	G	R	L	S	A	Y	I	NPA	NPA	109	plas
SINIP-2	267	28.27	7.78	0.66	6	W	V	A	R	F	S	A	Y	I	NPA	NPA	109	plas
SINIP-3	272	29.27	8.24	0.66	6	W	V	A	R	L	S	A	Y	I	NPA	NPA	109	plas
SINIP-4	277	29.66	9.16	0.43	6	W	V	A	R	F	S	A	Y	L	NPS	NPA	109	plas
SINIP-5	290	30.62	9.15	0.37	6	W	V	A	R	F	S	A	Y	M	NPA	NPA	109	plas
SINIP-6	345	37.22	8.30	0.44	6	W	I	A	R	F	S	A	Y	I	NPA	NPA	109	plas
SINIP-7	295	30.63	9.06	0.50	6	S	I	A	R	F	T	A	Y	L	NPS	NPV	108	plas
SINIP-8	260	27.82	8.91	0.68	6	W	V	A	R	F	S	A	F	V	NPA	NPA	112	vacu
SINIP-9	306	31.66	8.34	0.46	6	T	I	A	R	L	T	A	Y	L	NPA	NPV	108	plas
SINIP-10	231	25.45	9.62	0.45	4	/	S	G	R	F	S	A	Y	I	NPA	NPT	115	chlo
SINIP-11	137	14.86	6.70	0.75	3	W	/	/	/	L	/	/	/	/	NPA	/	/	vacu
SINIP-12	286	30.57	6.19	0.66	6	A	V	G	R	Y	S	A	Y	V	NPA	NPA	109	vacu
StLsi1	272	29.26	8.23	0.67	6	W	V	A	R	L	S	A	Y	I	NPA	NPA	109	plas
CsLsi1	288	30.55	9.40	0.33	6	G	S	G	R	L	T	A	Y	F	NPA	NPV	108	vacu
CsLsi6	261	27.62	6.05	0.45	6	C	S	G	R	M	S	A	Y	M	NPA	NPA	108	plas
HvLsi1	295	31.66	6.79	0.43	6	G	S	G	R	L	T	A	Y	F	NPA	NPA	108	plas
TaLsi1	295	31.67	6.79	0.42	6	G	S	G	R	L	T	A	Y	F	NPA	NPA	108	plas
SbLsi1	297	32.04	6.78	0.37	6	G	S	G	R	L	T	A	Y	F	NPA	NPA	108	plas
CmLsi1	297	32.04	6.78	0.37	6	G	S	G	R	L	T	A	Y	F	NPA	NPA	108	plas
ZmLsi1	295	31.76	6.79	0.35	6	G	S	G	R	L	T	A	Y	F	NPA	NPA	108	plas
OsLsi1	298	31.98	6.70	0.38	6	G	S	G	R	I	T	A	Y	F	NPA	NPA	108	plas
OsLsi6	298	31.83	7.00	0.49	6	G	S	G	R	L	T	A	Y	F	NPA	NPA	108	E.R
ZmLsi6	294	31.39	7.71	0.48	6	G	S	G	R	L	T	A	Y	F	NPA	NPA	108	plas
HvLsi6	300	32.25	7.75	0.44	6	G	S	G	R	L	T	A	Y	F	NPA	NPA	108	E.R

**Table 1.** Characteristics of tomato homologues of Si transporters, and analysis of their protein information, subcellular localization, conserved amino acid residues, conserved motifs and transmembrane domains (TMDs). *Len* length, *MW* molecular weight, *pI* isoelectric point, *GR* grand average of hydrophobicity, *TM* transmembrane domains, *ar/R* aromatic/arginine, *H2* helix2, *LE1* loop E1, *P1* position 1, *LB* loop B, *DB* distance, *LOC* localization.

(S or T), which are usually the sites of phosphorylation of proteins. This residue may be responsible for the post-transcriptional modification and regulation of protein activity. However, the amino acid composition of Position 5 differs from SINIP-1 to these already identified Si transporters. Position 5 is a nonaromatic residue in SINIP-1, and this residue is aromatic in these already identified Si transporters (Table 1). Furthermore, within tomato NIPs, residues at Position 1 are also different between SINIP-1 and members of tomato NIP I and NIP II subgroups. Residue at Position 1 is an aromatic residue in NIP I and NIP II members, and this residue is not aromatic in SINIP-1. Whether the variant residues at Position 1 and Position 5 in the Froger's positions affect transport activities of NIPs or not, which need to be experimentally proved in further studies.

The predicted results of physical and chemical properties of these Si transporters as well as members of tomato NIP I and NIP II are recorded in Table 1. Except SINIP-11 (137 AA and 14.86, respectively), the length of the encoded proteins mostly ranged from 231 to 345 amino acids, and their predicted molecular mass ranged from ~25.45 to 37.22 kDa. The isoelectric points of all 24 Si transporters ranged from 6.05 to 9.40, and the total average hydrophobicity of protein is between 0.33 and 0.75. The subcellular localization of these Si transporters and NIPs is predicted by WoLF online tool and recorded in Table 1. It is found that most of these proteins are located in plasma membrane, some are located in chloroplast membrane, vacuole membrane or endoplasmic reticulum. The difference between localizations may be related to the functional differences, and individual proteins may be involved in the transport of substances between plasma membrane and organelles.

**Tissue-specific expression analysis.** To explore the functions of tomato NIP genes as well as the two *SILsi2* genes, we firstly investigated their expression atlas based on the RNA-seq data derived from the TOMATO FUNCTIONAL GENOMICS DATABASE in 11 various tissues of tomato including whole root, hypocotyl, cotyledons, young leaves, mature leaves, vegetative meristems, young flower buds, flowers at anthesis (0 DPA), 10 days post anthesis fruit (10 DPA), 20 days post anthesis fruit (20 DPA) and ripening fruit (33 DPA) (Fig. 5). *SINIP-7* and *SINIP-2* exhibited high and organ-specific expression pattern. *SINIP-7* was highly expressed in all examined tissues, especially in root, indicating it may function in root. *SINIP-2* was preferentially expressed highly in anthesis flowers. *SINIP-5*, *SINIP-9* and *SILsi2-1* were constitutively expressed at a high level in nearly all tested tissues, and four genes (*SINIP-8*, *SINIP-10*, *SINIP-11* and *SINIP-12*) were expressed with very low transcript abundance. *SINIP-3* was preferentially expressed in anthesis flowers, while *SINIP-6* and *SILsi2-2* was spe-



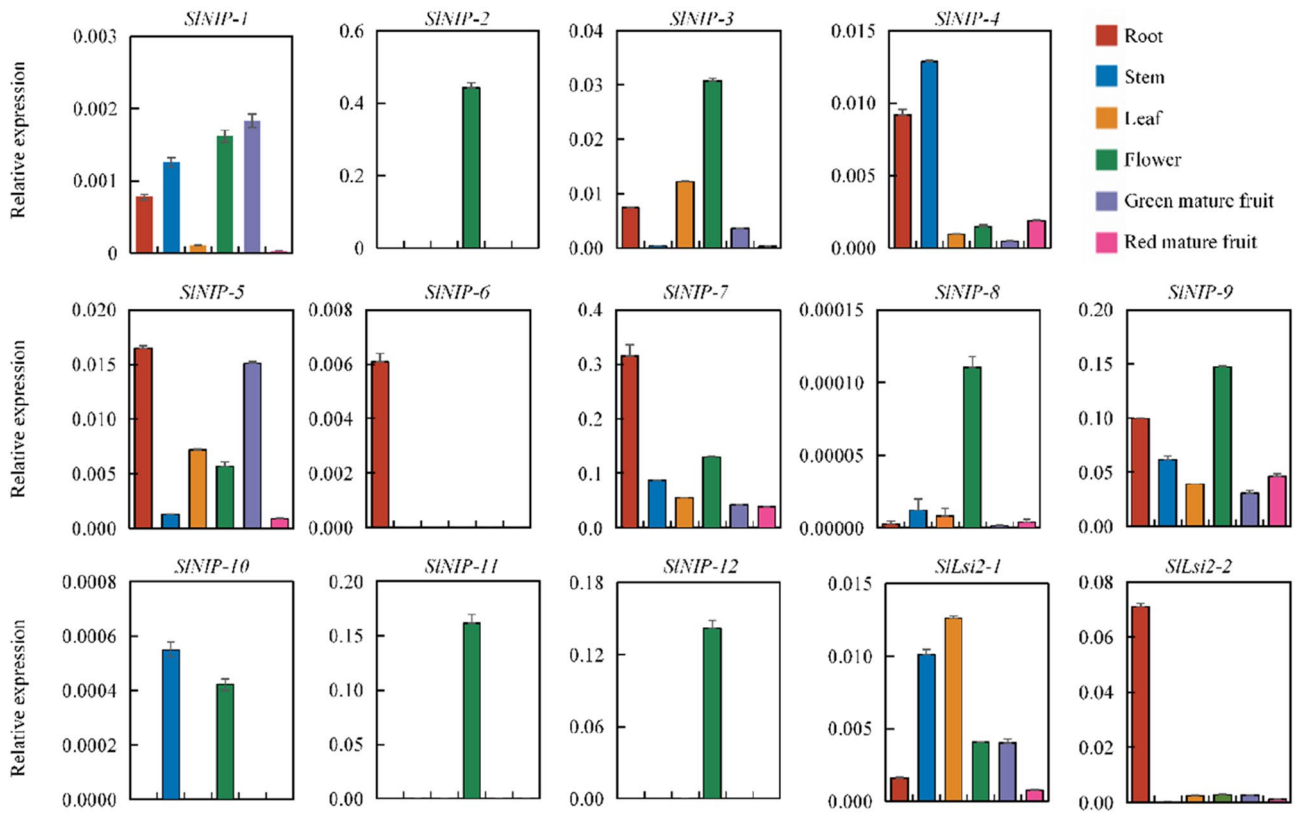
**Figure 5.** Expression profile of homologous genes of Si transporters in tomato in various organs available from public RNA-seq data. The expression data were log<sub>2</sub> transformation of RPKM (reads per kilobase of exon model per million mapped reads) with TBtools v1.0986988, and a cluster dendrogram is shown on the left of each heat map. Different colors represent differences in gene expression. Red indicates high concentrations, whereas low relative concentrations are blue. *Root* whole root, *Hypo* hypocotyl, *Cotyl* Cotyledon, *YL* young leaf, *ML* mature leaf, *MERI* vegetative meristem, *YFB* young flower bud, *0 DPA* flower at anthesis, *10 DPA* fruit at 10 DPA, *20 DPA* fruit at 20 DPA, *33 DPA* fruit at 33 DPA.

cially expressed in root. A similar expression pattern was observed on *SINIP-4* and *SINIP-5* genes, which showed relatively higher expression levels in root, hypocotyl, cotyledons, and fruit (20 DPA), but a lower level of expression in ripening fruit. *SINIP-1* showed relatively higher expression levels in root, hypocotyl, cotyledons, anthesis flowers, immature fruit, but relatively lower expression levels in young leaves, mature leaves, vegetative meristems, young flower buds and ripening fruit, indicating *SINIP-1* may function during fruit development (Fig. 5).

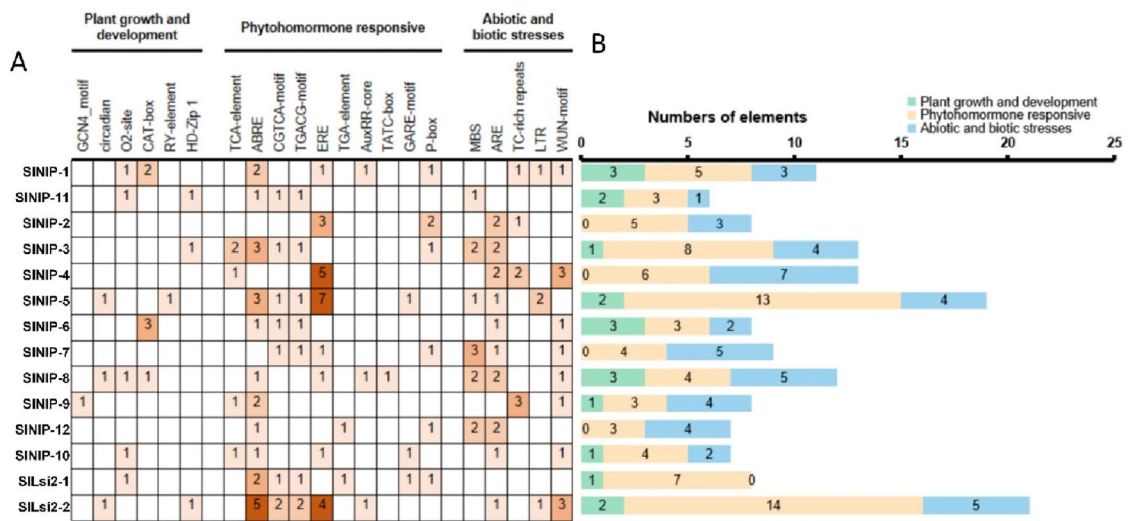
To further explore the possible functions of tomato NIP genes, organ-specific expression profiles of these genes were further verified by RT-qPCR. We used 6 different tissues (root, stem, leaf, flower, mature green fruit, and red mature fruit). As shown in Fig. 6, expression patterns of most of these genes were consistent with the RNA-seq data in Fig. 5. Consistent with previous report, the already identified Si influx transporter *SINIP-1* was mainly expressed in root relative to leaf, and Si efflux transporter *SILsi2-2* was specific to root (Fig. 6<sup>30</sup>). Inconsistent with previous report<sup>30</sup>, another efflux transporter *SILsi2-1* was expressed mainly in leaf but not in root, it may be due to *SILsi2-1* was not a consecutive expression gene but an induced gene. Six genes (*SINIP-2*, *SINIP-3*, *SINIP-8*, *SINIP-9*, *SINIP-11* and *SINIP-12*) had similar expression patterns and demonstrated a high expression in flower. *SINIP-4* was expressed strongly in root and stem, while *SINIP-10* was expressed intensely in stem and flower. *SINIP-5* was expressed significantly in root and green mature fruit, weakly in stem and red mature fruit, and moderately in leaf and flower, whereas, *SILsi2-1* was expressed highly in stem and leaf, weakly in root and red mature fruit, and moderately in flower and green mature fruit. *SINIP-6* like *SILsi2-2* displayed root-specific expression patterns. *SINIP-7* was constitutively expressed at a high level in almost all tested tissues, and showed relatively higher expression levels in root. Above all, different members of the tomato NIP genes showed distinct expression patterns, indicating these genes may function in many aspects of tomato growth and development.

**Promoter analysis.** Specific cis-acting elements can combine with particular transcription factors to regulate downstream gene transcription and expression. In order to determine the promoter cis-acting elements of 14 tomato Si transporter homologous proteins, the 2000 bp sequences upstream of the start site (ATG) of these genes were downloaded as putative promoters from SGN database. The promoter sequences were analyzed by using PlantCARE online tool, and the identified cis-acting elements were sorted into three groups related to plant growth and development, hormone response and biotic/abiotic stresses (Fig. 7). Cis-acting elements related to plant growth and development included endosperm expression element (GCN4\_motif), circadian rhythm regulation element (circadian), zein metabolism regulation element (O<sub>2</sub>-site), meristem expression element (CAT-box), seed-specific regulation element (RY-element) and mesophyll cell differentiation element (HD-Zip1). There are only 0–3 cis-acting elements of this group in promoters of all 14 tomato Si transporter homologous genes, which indicates that these genes may be less induced by signaling from the specific growth and development stage of plants. Cis-acting elements belonging to hormone response included salicylic acid responsive element (TCA-element), abscisic acid (ABA) responsive element, methyl jasmonate responsive element (CGTCA-motif, TGACG-motif), ethylene responsive element (ERE), auxin responsive element (TGA-element, AuxRR-core) and gibberellin responsive element (TATC-box, GARE-motif, P-Box), which has 3–14 cis-acting elements of this group in the promoters of 14 tomato Si transporter homologues, indicating that these genes are possibly sensitive to plant hormones. Eleven tomato Si transporter homologues except *SINIP-2*, *SINIP-4* and *SINIP-7* all have ABA responsive elements, while eight tomato Si transporter homologues except *SINIP-3*, *SINIP-6*, *SINIP-9*, *SINIP-11*, *SINIP-12* and *SILsi2-1* all have ethylene responsive element (ERE). ABA and

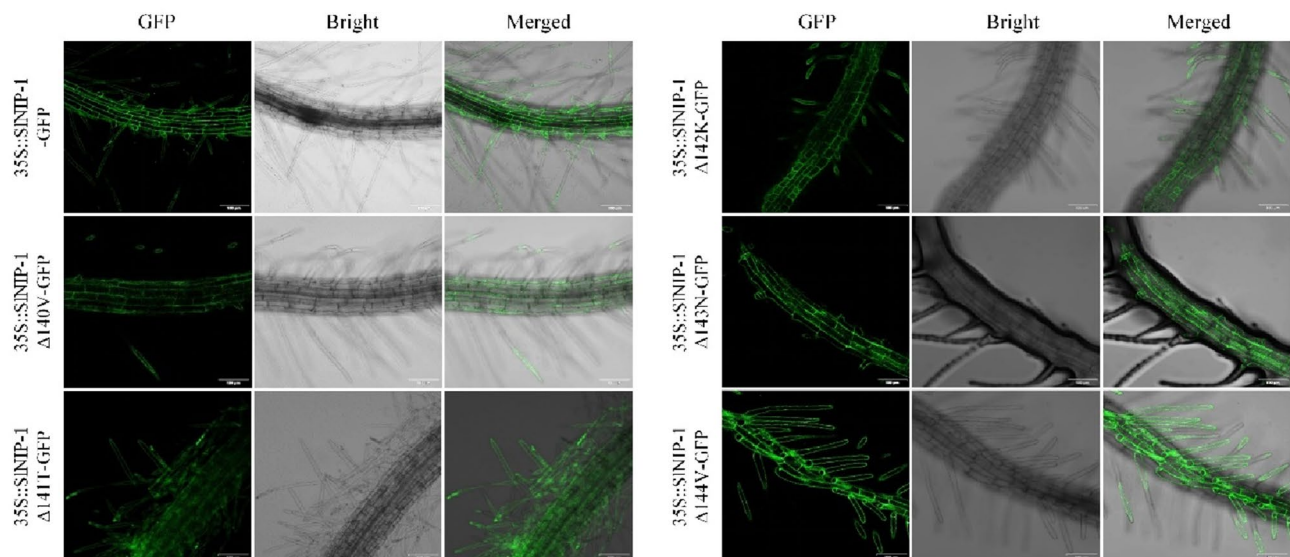




**Figure 6.** Expression patterns of homologous genes of Si transporters in tomato in different tissues were examined using qRT-PCR. Tomato *ACTIN2* gene was used as an internal standard. Values are means  $\pm$  SD (n = 3). Gray represents the roots of tomato; yellow represents the stems of tomato; blue represents the leaves of tomato; red represents the flowers of tomato; green represents the green mature fruits of tomato; pink represents the red mature fruits of tomato.



**Figure 7.** Prediction of cis-acting elements in promoter regions of Si transporter homologues in tomato. (A) The different colors and numbers of the grid indicated the numbers of each predicted element in these Si transporter homologues promoter region. (B) The different colored histogram represented the sum of the cis-acting elements in each category.



**Figure 8.** Subcellular localization of SINIP-1 protein and its mutations of site-specific deletion. The roots of seven-day-old seedlings were used and the fluorescence were monitored. Bars 100  $\mu\text{m}$ . Vectors of *35::SINIP-1-GFP*, *35::SINIP-1 $\Delta$ 140V-GFP*, *35::SINIP-1 $\Delta$ 141T-GFP*, *35::SINIP-1 $\Delta$ 142K-GFP*, *35::SINIP-1 $\Delta$ 143N-GFP*, and *35::SINIP-1 $\Delta$ 144V-GFP* were introduced into *Arabidopsis thaliana* ecotype Columbia mediated by *Agrobacterium tumefaciens* strain GV3101. T<sub>1</sub> transgenic plants were selected by genotyping.

ethylene play vital roles in plant growth and development as well as in mediating plant response to a wide range of stresses. It should be noted that tomato Si transporter homologues may be regulated by ABA or ethylene and participate ABA or ethylene-mediated plant development and stress responsiveness. Cis-acting elements related to biotic/abiotic stress included drought response element (MBS), anaerobic induction response element (ARE), defense and stress response element (TC-rich repeats), low temperature response element (LTR) and mechanical injury response element (WUN-motif). There are 1–7 cis-acting elements of this group in the promoters of 14 tomato Si transporter homologues. Among them, three WUN-motifs, MBS, TC-rich repeats were found in *SINIP-4*, *SILsi2-2*, *SINIP-7* and *SINIP-9* promoters, respectively. The cis-regulatory elements identified in the promoters of tomato Si transporter homologues suggested that they may play an important role in hormone signaling and various stress response.

**Subcellular localization of SINIP-1 and its mutated proteins.** Plasma membrane location is required for the uptake of Si by Lsi1 transporters. Functional Lsi1 Si influx transporters are unexceptionally localized in the plasma membrane, such as OsLsi1<sup>13</sup>, HvLsi1<sup>33</sup>, ZmLsi1<sup>34</sup>, TaLsi1<sup>22</sup>, SILsi1<sup>29</sup>, CsLsi1<sup>35</sup>, and NsLsi1<sup>41</sup>. Subcellular localization can largely affect the Si transport activities. CmLsi1 from the bloomless rootstock that was localized at the plasma membrane have transport activity, whereas the one from the bloomless rootstock was localized at the endoplasmic reticulum and had no transport activity for Si. In order to verify the subcellular localization of tomato SINIP-1, we cloned the 849 bp CDS sequence of *SINIP-1*, and introduced it into the vector *pGWB5* to obtain the GFP fusion vector *SINIP-1-pGWB5* (*35::SINIP-1-GFP*). The vector was transformed into *Agrobacterium tumefaciens* GV3101 to generate transgenic *Arabidopsis* plants by floral dip method. T<sub>1</sub> transgenic *Arabidopsis* plants were used for observing GFP fluorescence signal. Consistent with WoLF's prediction in Table 1 and previous report<sup>29</sup>, SINIP-1 protein was localized in plasma membrane (Fig. 8). A 108-AA spacing between NPA domains is essential to Si influx transport activity for SILsi1<sup>29</sup>. Predicted by PROVEAN, we found five site deletions (140 V, 141 T, 142 K, 143 N and 144 V) result in neutral effects. Then, the localization of variant SINIP-1 mutants of single amino acid deletion was confirmed by transgenic plants. As the results showed that, like SINIP-1, these five mutated SINIP-1 proteins (SINIP-1 $\Delta$ 140V, SINIP-1 $\Delta$ 141T, SINIP-1 $\Delta$ 142K, SINIP-1 $\Delta$ 143N, and SINIP-1 $\Delta$ 144V) with a spacing of 108 AA were also located on the plasma membrane (Fig. 8). In 2015, Deshmukh et al. had reported that deletion 140 V in SINIP-1 exhibited Si permeability relative to the native SINIP-1 protein<sup>29</sup>. Thus, the function of the other four SINIP-1 mutants need to be tested.

## Discussion

Tomato Si transporter homologues SINIP-1 to SINIP-12 belong to NIP subfamily (Figs. 1, 3). Six of the twelve proteins with ar/R filter consisting of Trp (H2), Val (H5), Ala (LE1), and Arg (LE2), WVAR, as well as SINIP-6 with WIAR, belong to the NIP I subgroup, whereas SINIP-7 with SIAR, SINIP-9 with TIAR, and SINIP-12 with AVGR were classified into NIP II group<sup>42,43</sup>. Additionally, SINIP-1 with GSGR belong to NIP III group<sup>28</sup>. Compared to NIP III subgroups, which have small amino acid residues Gly in H2 and Ser in H5; members of NIP I subgroups have two bulky residues Trp and Val in H2 and H5 positions, respectively<sup>23</sup>. So, considering the composition of ar/R filter, we propose that most members of tomato NIPs possessing WVAR (WIAR in SINIP-6) do not have transport activities for silicic acid, but may transport small molecules of water, glycerol

and lactic acid as members in NIP I. To support the propose, one study showed that replacement of AIGR by WIGR endows AtNIP6;1 with water transport activity, suggesting the possible water-transport activity as well as glycerol transport of SINIP-6 which was presented by WIAR of the ar/R region.

The Arabidopsis NIP II subgroup has three member genes: AtNIP5;1, AtNIP6;1, and AtNIP7;1<sup>23</sup>, and each of them has an ortholog in tomato, represented by SINIP-7, SINIP-9 and SINIP-12, respectively<sup>43</sup>. Current study further confirmed the remarkable orthologous relationships between AtNIP5;1, AtNIP6;1, AtNIP7;1 and SINIP-7, SINIP-9, SINIP-12. Firstly, structural similarity is within pore determinant regions, such as SINIP-7 and AtNIP5;1 share the a/R filter of SIAR; SINIP-9 and AtNIP6;1 share the a/R filter of TIAR; SINIP-12 together with AtNIP7;1 share the a/R filter of AVGR (Table 1). Secondly, similar tissue expression profiles, RT-qPCR combined with RNA-seq data showed that SINIP-7 was mainly expressed in the roots, SINIP-9 was constitutively expressed in all examined tissues with high expression level in stem, and SINIP-12 was specifically expressed in flower (Fig. 5), which were similar to previous studies<sup>43–45</sup>. Thirdly, similar to previous observations with AtNIP5;1<sup>45</sup>, AtNIP6;1<sup>44</sup> and AtNIP7;1<sup>43</sup>, SINIP-7 and SINIP-9 were predicted to be localized on plasma membrane, and SINIP-9 was predicted to be localized on vacuole membrane (Table 1). As AtNIP5;1, AtNIP6;1, and AtNIP7;1 were reported to mediate boron uptake and contribute to the development of root, rosette leaves, inflorescences and pollen grain<sup>27,43,44</sup>. In summary, the results suggested SINIP-7, SINIP-9 and SINIP-12 might act as major channel proteins mediating boric acid transport. It will be important to investigate the substrate specificity of the SINIP-7, SINIP-9 and SINIP-12 and the possible involvement of SINIP-7, SINIP-9 and SINIP-12 in plant growth and development.

In addition to ar/R selectivity filter, NPA motif and inter-NPA distance can also affect the Si permeability of Lsi1 channels. It has been reported that two NPA motifs localized at loop B and loop E form part surface of the narrow aqueous pore<sup>46</sup>. Residues composed of NPA motifs influence the size of constriction filter, and then the specificity of substrates. For example, the pore diameter of the NPA regions of AtNIP6;1 is narrower than that of Nodulin 26, which may result from a larger Val substitution for the small Ala in the second NPA motif (loop E) of AtNIP6;1 compared with Nodulin 26<sup>26</sup>. Most members of the tomato NIPs have two conserved NPA motifs, while the first NPA motif in SINIP-4 and SINIP-7 was changed to NPS. In addition, the second NPA motif in SINIP-7 was also replaced by NPV. Moreover, the second motif in SINIP-9 and SINIP-10 was changed to NPV and NPT, respectively. However, it seems that the second NPA motif is not a crucial factor for the substrate specificity of NIP proteins. In the case of AtNIP6;1, a substitution of Ala for Val in the second NPA motif has little effect on the transport selectivity of AtNIP6;1<sup>26</sup>.

An inter-NPA distance of 108-AA is a common feature that is critical for the function of Si transporters. SINIP-1/SILsi1 owned an ar/R filter composing of GSGR belonging to the NIP III group and might be a Si transporter, but 109 AA in the inter-NPA region of SINIP-1 may cause the loss function of Si transport activities. However, Si permeability of SILsi1 is still controversial<sup>29,30</sup>. In addition to the precise space between the NPA domains, other molecular determinants can affect the function of Lsi1 transporters. NsLsi1 from tobacco possessing typical molecular signatures, a GSGR selectivity filter and a 108 AA inter-NPA space, were found to be Si-impermeable<sup>41</sup>. Further investigation showed that Si transport activities of NsLsi1 can be compromised by P125F substitution. P125F substitution increased plasma membrane localization of NsLsi1<sup>P125F</sup> compared to NsLsi1<sup>WT</sup>, thus enhanced the transport capacity<sup>41</sup>. It was also reported that analysis of boric acid permeability of AtNIP7;1 by expressing AtNIP7;1 in *Xenopus laevis* oocytes and proteoliposomes yielded contrasting results, explaining by some factors influencing the pore structures of the AtNIP7;1 changed in this two systems<sup>43</sup>. As for the case of SINIP-1, other than the molecular characteristics of protein sequences, determining factors, such as biochemical properties as well as posttranslational modification, which affect the function of SINIP-1 may be changed due to different analysis systems and methods. Besides, deletion 140 V in SILsi1/SINIP-1 restored its Si permeability<sup>29</sup>. In the future, the activity of SINIP-1Δ141T, SINIP-1Δ142K, SINIP-1Δ143N, and SINIP-1Δ144V are worth testing. Finally, definition of the in vivo transport properties of SINIP-1 will require further detailed investigation.

Tissue-specific expression pattern reflects the function of genes to some extent. We conducted RNA-seq and RT-qPCR analysis, discovering some interesting trends and expression patterns of 14 tomato Si transporter homologues (Figs. 5, 6). SINIP-1 and SILsi2-2 were mainly expressed in root relative to leaf, which is consistent with their function reported in previous study<sup>30</sup>. SILsi2-1 was found mainly expressed in leaf but not root, which did not match previous results<sup>30</sup>. This discrepancy may reflect that SILsi2-1 was not expressed constantly but can be induced by some stimulation. The high expression pattern in root was also found in SINIP-4, SINIP-5, and SINIP-6, implying their roles in root. Indeed, AtNIP2;1 from *Arabidopsis* was a root-specific expression gene, and it was proved to be a lactic acid efflux channel and is necessary for an optimum response to low oxygen stress<sup>47</sup>. Thus, considering SINIP-4, SINIP-5 and SINIP-6 both belong to the NIP I family and with similar root-specific expression as AtNIP2;1, we propose that SINIP-4, SINIP-5 and SINIP-6 may be involved in abiotic stress response in root.

SINIP-2, SINIP-8, and SINIP-11 exhibited an almost undetectable low expression level in most examined tissues but show high expression levels in flower, suggesting these genes may play specific roles in flower. Similarly, SINIP-3 and SINIP-10 were also mainly expressed in flower. Polygenetic tree indicated that SINIP-2, SINIP-3 and SINIP-11 were highly homologous with AtNIP4;1 and AtNIP4;2. It has been reported that AtNIP4;1 and AtNIP4;2 were expressed in flower specific to pollen grains and pollen tubes, and characterized as important regulators for the proper pollen development, pollen germination, and pollen tube growth<sup>48</sup>. Similar to AtNIP4;1 and AtNIP4;2, SINIP-2, SINIP-3 and SINIP-11 may participate in the male gametophyte development.

## Conclusions

There are 43 aquaporins including 8 NIP I proteins, 3 NIP II proteins, and 1 NIP III protein in tomato. SINIP-1 was the reported Si influx transporters SILsi1. Mutated SINIP-1 proteins of single amino acid deletion—SINIP-1 $\Delta$ 141T, SINIP-1 $\Delta$ 142K, SINIP-1 $\Delta$ 143N, and SINIP-1 $\Delta$ 144V—resembled the native SINIP-1 showing plasma membrane localization. Though polygenetic tree analysis, conserved structure analysis, and gene expression patterns analysis, we predict SINIP-7, SINIP-9, SINIP-12 may be boric acid facilitators, and SINIP-2, SINIP-3 and SINIP-11 may be involved in the pollen development. Considering the important biological functions of NIP subfamily of plant aquaporins, further analysis of the functions of these genes in specific tissue is deserved. In addition, we also propose that functional identification of SINIP-1, SILsi2-1 and SILsi2-2 should be obtained through transgenic plants and furthermore experimental studies.

## Materials and methods

**Screening and identification of Lsi1 and Lsi6 homologues and Lsi2 homologues in tomato.** The amino acid sequences of Si transporters Lsi1, Lsi2 and Lsi6 of rice and cucumber were obtained from NCBI database (<https://www.ncbi.nlm.nih.gov/>). Sequences of tomato were searched by the BLAST program provided by tomato genome database SGN (<https://solgenomics.net/>), and then the sequences of retrieved homologous proteins with E value less than  $e^{-6}$  was selected. The amino acid sequences of barley (*Hordeum vulgare* L.), wheat (*Triticum aestivum* L.), corn (*Zea mays* L.), sorghum (*Sorghum bicolor* L.), pumpkin (*Cucurbita moschata* L.) and potato (*Solanum tuberosum* L.) were downloaded from NCBI database.

**Phylogenetic tree construction and chromosome location.** Sequence alignments of Si transporters of tomato and other species were carried out by ClustalX2 software (<http://www.clustal.org/>)<sup>49</sup>, and sequence alignment and evolutionary tree construction were further carried out by MEGA7.0 software (<https://www.megasoftware.net/>)<sup>50</sup>. Neighbor-joining method was selected as the evolutionary tree construction method with 1000 bootstrap replications. Finally, the phylogenetic tree was illustrated using Interactive Tree of Life (IToL, <http://itol.embl.de>) and Notepad++ v7.8.8 software (<https://notepad-plus-plus.org/downloads/v7.8.8/>). The chromosomal locations of tomato Lsi1, Lsi2 and Lsi6 homologous genes were searched in tomato genome database SGN, and the distribution of gene location on chromosome was drawn by MapInspect 1.0 software (<https://mapinspect.software.informer.com/>).

**Analysis of exon–intron structure and conserved motifs.** The gene sequences of tomato Si transporter homologous proteins were searched from tomato genome database SGN, and the exon and intron of these genes were found on GSDS website (<http://gsds.cbi.pku.edu.cn/>); the motifs of these proteins was searched by MEME website (<https://meme-suite.org/meme/>). The method was set to Any Number of Repetitions, and the number of motifs searched was set to 10. The transmembrane domains were analyzed by TMHMM online tool (<http://www.cbs.dtu.dk/services/TMHMM/>) prediction; subcellular localizations were predicted by WoLF PSORT Tool (<https://wolfsort.hgc.jp/>). The basic physical and chemical properties of protein were analyzed by ProtParam online tool (<http://web.expasy.org/protparam/>); the positions of ar/R selectivity filter and Froger's residues were marked after sequence alignments.

**Evolutionary analysis by maximum likelihood method.** The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model. The tree with the highest log likelihood (−14,866.98) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. This analysis involved 33 amino acid sequences. There were a total of 474 positions in the final dataset. Evolutionary analyses were conducted in MEGA 7.0.

**Tissue-specific expression analysis.** The expression patterns analysis of tomato Si transporter homologue genes in different tissues of different development stages of tomato was carried out according to RNA-seq data from Tomato Functional Genomics Database (TFGD, <http://ted.bti.cornell.edu/cgi-bin/TFGD/digital/home.cgi>). Transcriptome analysis of 11 tissues in wild species *S. pimpinellifolium*, LA1589 using Illumina RNA-seq. The tissues include: newly developed leaves around 5 mm long, mature green leaflets, flower buds 10 days before anthesis or younger, flowers at anthesis, fruit of 10 days post anthesis (DPA), 20 DPA fruit, and breaker stage ripening fruit. RPKM values are an average of 4 replicates. Gene expression data of RPKM values were log2 transformed before analysis, and heatmaps were generated by TBtools v1.0986988 software<sup>51</sup>.

**Plant materials and growth conditions.** Tomato (*Solanum lycopersicum*) cv. Ailsa Craig (AC) was used as the wild type in this study. Seeds of tomato cultivar AC (Accession LA2838A) were obtained from the Tomato Genetics Resources Center at the University of California, Davis (<https://tgrc.ucdavis.edu/>). Tomato seeds were placed on moistened filter paper for 48 h for germination. Tomato seedlings were transferred to growth chambers and maintained under a long-day photoperiod (16 h of light/8 h of dark) with a white light intensity of 200 mmol photons  $m^{-2} s^{-1}$  at 25 °C during the subjective day and at 18 °C during the subjective night.

*Arabidopsis thaliana* ecotype Columbia was used as the wild type in this study. *Arabidopsis* plants were grown in Murashige and Skoog (MS) media at 22 °C under a long-day (16 h light/8 h dark) photoperiod (light intensity of 120 mmol photons  $m^{-2} s^{-1}$ ).

The plant materials were well used for research project and comply with relevant institutional, national, and international guidelines and legislation.

**RNA extraction and quantitative RT-PCR (qRT-PCR) analysis.** TRIzol reagent (Invitrogen) was used for RNA extraction. For quantitative real-time polymerase chain reaction (qRT-PCR) analysis, one microgram of DNA-free RNA was transcribed into first strand cDNA by Prime-Script RT Master Mix (TaKaRa). The qRT-PCR was carried out with the UVP ChemStudio (analyticjena) using TB Green Premix Ex Taq (TaKaRa). The reaction conditions were 95 °C for 30 s, and 40 cycles at 95 °C for 5 s, 60 °C for 30 s. Expression levels of target genes were normalized relative to *ACTIN2* gene. Primers used to quantify gene expression levels are listed in Supplementary Table S4. Each reaction was performed with three biological replicates.

**Cis-elements analysis.** Regions 2000 bp upstream of the start codon (ATG) of each gene were downloaded in the tomato genome database SGN as the predicted promoter sequences. Cis-acting elements of the promoters were predicted by PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>).

**Plasmid construction and plant transformation.** To generate the *35S::SINIP-1-GFP* construct, the coding sequence without termination codon of *SINIP-1* was cloned into pENTR and recombined with the binary vector *PGWB5* (35S promoter, C-GFP). A similar approach was used to generate the *35S::SINIP-1Δ140V-GFP* (35S promoter, C-GFP), *35S::SINIP-1Δ141T-GFP*, *35S::SINIP-1Δ142K-GFP*, *35S::SINIP-1Δ143N-GFP*, *35S::SINIP-1Δ144V-GFP* constructs. All primers used for DNA construct generation are listed in Supplemental Table S4. The above constructs were then transformed into *Agrobacterium tumefaciens* strain GV3101, which was used for transformation of *Arabidopsis* plants via the floral dip method. T<sub>1</sub> transgenic plants were selected based on their resistance to hygromycin.

**Microscopy observations.** For subcellular localization analysis, roots of 10-day-old T<sub>1</sub> transgenic plants with corresponding constructs were observed and imaged under a laser-scanning confocal microscope (Olympus fluoview FV3000). For imaging GFP, the 488-nm lines of the laser were used for excitation, and emission was detected at 510 nm.

## Data availability

The amino acid sequences analyzed in the current study are available in the Sol Genomics Network repository (<https://solgenomics.net/>), the NCBI database (<https://www.ncbi.nlm.nih.gov/>), the Arabidopsis Information Resource (TAIR) (<http://www.arabidopsis.org/>) and the rice genome annotation database (<http://rice.uga.edu/>), respectively. The accession numbers are included in the Supplementary Tables.

Received: 22 February 2022; Accepted: 23 May 2022

Published online: 02 June 2022

## References

- Jugdaohsingh, R. Silicon and bone health. *J. Nutr. Health Aging* **11**, 99–110 (2007).
- Hans Wedepohl, K. The composition of the continental crust. *Geochim. Cosmochim. Acta* **59**, 1217–1232 (1995).
- Epstein, E. The anomaly of silicon in plant biology. *Proc. Natl. Acad. Sci. USA* **91**, 7–11 (1994).
- Ma, J. F. Role of silicon in enhancing the resistance of plants to biotic and abiotic stresses. *Soil Sci. Plant Nutr.* **50**, 11–18 (2004).
- Zhu, Y. & Gong, H. Beneficial effects of silicon on salt and drought tolerance in plants. *Agron. Sustain. Dev.* **34**, 455–472 (2014).
- Adrees, M. *et al.* Mechanisms of silicon-mediated alleviation of heavy metal toxicity in plants: A review. *Ecotoxicol. Environ. Saf.* **119**, 186–197 (2015).
- Jones, L. H. P. & Handreck, K. A. In *Advances in Agronomy* (ed. Norman, A. G.) 107–149 (Academic Press, 1967).
- Abascal, F., Irisarri, I. & Zardoya, R. Diversity and evolution of membrane intrinsic proteins. *Biochim. Biophys. Acta* **1840**, 1468–1481 (2014).
- Ma, J. F., Miyake, Y. & Takahashi, E. In *Studies in Plant Science* (eds Datnoff, L. E. *et al.*) 17–39 (Elsevier, 2001).
- Nikolic, M., Nikolic, N., Liang, Y., Kirkby, E. A. & Römheld, V. Germanium-68 as an adequate tracer for silicon transport in plants. Characterization of silicon uptake in different crop species. *Plant Physiol.* **143**, 495–503 (2006).
- Ma, J. F. & Yamaji, N. Functions and transport of silicon in plants. *Cell Mol. Life Sci.* **65**, 3049–3057 (2008).
- Mitani, N. & Ma, J. F. Uptake system of silicon in different plant species. *J. Exp. Bot.* **56**, 1255–1261 (2005).
- Ma, J. F. *et al.* A silicon transporter in rice. *Nature* **440**, 688–691 (2006).
- Ma, J. F. *et al.* An efflux transporter of silicon in rice. *Nature* **448**, 209–212 (2007).
- Yamaji, N., Mitani, N. & Ma, J. F. A transporter regulating silicon distribution in rice shoots. *Plant Cell* **20**, 1381–1389 (2008).
- Yamaji, N. & Ma, J. F. A transporter at the node responsible for intervacular transfer of silicon in rice. *Plant Cell* **21**, 2878–2883 (2009).
- Gomes, D. *et al.* Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. *Biochim. Biophys. Acta Biomembr.* **1788**, 1213–1228 (2009).
- Abascal, F., Irisarri, I. & Zardoya, R. Diversity and evolution of membrane intrinsic proteins. *Biochim. Biophys. Acta* **1840**, 1468–1481 (2005).
- Johanson, U. & Gustavsson, S. A new subfamily of major intrinsic proteins in plants. *Mol. Biol. Evol.* **19**, 456–461 (2002).
- Bienert, G. P., Bienert, M. D., Jahn, T. P., Boutry, M. & Chaumont, F. *Solanaceae* XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates. *Plant J.* **66**, 306–317 (2011).
- Danielson, J. A. & Johanson, U. Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. *BMC Plant Biol.* **8**, 45 (2008).
- Mitani-Ueno, N., Yamaji, N., Zhao, F. J. & Ma, J. F. The aromatic/arginine selectivity filter of NIP aquaporins plays a critical role in substrate selectivity for silicon, boron, and arsenic. *J. Exp. Bot.* **62**, 4391–4398 (2011).
- Wallace, I. S. & Roberts, D. M. Homology modeling of representative subfamilies of Arabidopsis major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter. *Plant Physiol.* **135**, 1059–1068 (2004).

24. Dean, R. M., Rivers, R. L., Zeidel, M. L. & Roberts, D. M. Purification and functional reconstitution of soybean nodulin 26. An aquaporin with water and glycerol transport properties. *Biochemis* **38**, 347–353 (1999).
25. Choi, W. G. & Roberts, D. M. Arabidopsis NIP2;1, a major intrinsic protein transporter of lactic acid induced by anoxic stress. *J. Biol. Chem.* **282**, 24209–24218 (2006).
26. Wallace, I. S. & Roberts, D. M. Distinct transport selectivity of two structural subclasses of the nodulin-like intrinsic protein family of plant aquaglyceroporin channels. *Biochemis* **44**, 16826–16834 (2005).
27. Takano, J. *et al.* The Arabidopsis major intrinsic protein NIP5;1 is essential for efficient boron uptake and plant development under boron limitation. *Plant Cell* **18**, 1498–1509 (2006).
28. Saitoh, Y. A. O. *et al.* Structural basis for high selectivity of a rice silicon channel Lsi1. *Nat. Commun.* **12**, 6236 (2021).
29. Deshmukh, R. A. O. *et al.* A precise spacing between the NPA domains of aquaporins is essential for silicon permeability in plants. *Plant J.* **83**, 489–500 (2015).
30. Sun, H. *et al.* Tomato roots have a functional silicon influx transporter but not a functional silicon efflux transporter. *Plant Cell Environ.* **43**, 732–744 (2020).
31. Wang, M. *et al.* Functions of silicon in plant drought stress responses. *Hortic. Res.* **8**, 254 (2021).
32. Routray, P. *et al.* Nodulin intrinsic protein 7;1 is a tapetal boric acid channel involved in pollen cell wall formation. *Plant Physiol.* **178**, 1269–1283 (2018).
33. Reuscher, S. *et al.* Genome-wide identification and expression analysis of aquaporins in tomato. *PLoS One* **8**, e79052 (2008).
34. Chiba, Y., Mitani, N., Yamaji, N. & Ma, J. F. HvLsi1 is a silicon influx transporter in barley. *Plant J.* **57**, 810–818 (2009).
35. Mitani, N., Yamaji, N. & Ma, J. F. Identification of maize silicon influx transporters. *Plant Cell Physiol.* **50**, 5–12 (2009).
36. Sun, H. *et al.* Isolation and functional characterization of CsLsi1, a silicon transporter gene in *Cucumis sativus*. *Physiol. Plant* **159**, 201–214 (2017).
37. Johanson, U. *et al.* The complete set of genes encoding major intrinsic proteins in Arabidopsis provides a framework for a new nomenclature for major intrinsic proteins in plants. *Plant Physiol.* **126**, 1358–1369 (2001).
38. Forrest, K. L. & Bhav, M. Major intrinsic proteins (MIPs) in plants: A complex gene family with major impacts on plant phenotype. *Funct. Integr. Genom.* **7**, 263–289 (2007).
39. Wallace, I. S., Choi, W. G. & Roberts, D. M. The structure, function and regulation of the nodulin 26-like intrinsic protein family of plant aquaglyceroporins. *Biochim. Biophys. Acta* **1758**, 1165–1175 (2006).
40. Froger, A., Tallur, B., Thomas, D. & Delamarque, C. Prediction of functional residues in water channels and related proteins. *Protein Sci.* **7**, 1458–1468 (1998).
41. Montpetit, J. *et al.* Cloning, functional characterization and heterologous expression of TaLsi1, a wheat silicon transporter gene. *Plant Mol. Biol.* **79**, 35–46 (2012).
42. Coskun, D. *et al.* Si permeability of a deficient Lsi1 aquaporin in tobacco can be enhanced through a conserved residue substitution. *Plant Direct* **3**, e00163 (2019).
43. Mitani, N., Yamaji, N. & Ma, J. F. Characterization of substrate specificity of a rice silicon transporter, Lsi1. *Pflug. Arch. Eur. J. Phys.* **456**, 679–686 (2008).
44. Tanaka, M., Wallace, I. S., Takano, J., Roberts, D. M. & Fujiwara, T. NIP6;1 is a boric acid channel for preferential transport of boron to growing shoot tissues in Arabidopsis. *Plant Cell* **20**, 2860–2875 (2008).
45. Wang, S. A. O. *et al.* Polar localization of the NIP5;1 boric acid channel is maintained by endocytosis and facilitates boron transport in Arabidopsis roots. *Plant Cell* **29**, 824–842 (2017).
46. Murata, K. *et al.* Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599–605 (2000).
47. Beamer, Z. A. O. *et al.* Aquaporin family lactic acid channel NIP2;1 promotes plant survival under low oxygen stress in Arabidopsis. *Plant Physiol.* **187**, 2262–2278 (2021).
48. Di Giorgio, J. A. O. *et al.* Pollen-specific aquaporins NIP4;1 and NIP4;2 are required for pollen development and pollination in *Arabidopsis thaliana*. *Plant Cell* **28**, 1053–1077 (2016).
49. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
50. Kumar, S. *et al.* MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
51. Chen, C. *et al.* TBtools—an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **13**, 1194–1202 (2020).

## Acknowledgements

The study was supported by the National Key Research and Development Program of China (2018YFD1000800), the Natural Science Foundation of Zhejiang Province (LZ20C150001), the National Natural Science Foundation of China (32102402) and the Scientific Research and Development Foundation of Zhejiang A & F University (2019FR045).

## Author contributions

Z.J.Z. and Y.Y.L. designed the experiments. Y.X.Z., Y.Y.L. and S.H.F. performed the experiments. Y.X.Z. and Y.Y.L. analyzed the data. Y.Y.L. wrote the original draft of the manuscript. Y.M.X., Y.H., Z.J.Z. and Y.Y.L. revised article. All authors read and approved of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-13195-0>.

**Correspondence** and requests for materials should be addressed to Z.Z. or Y.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022